

# Existence of Codes With Constant PMEPR and Related Design

Masoud Sharif, *Student Member, IEEE*, and Babak Hassibi

**Abstract**—Recently, several coding methods have been proposed to reduce the high peak-to-mean envelope ratio (PMEPR) of multicarrier signals. It has also been shown that with probability one, the PMEPR of any random codeword chosen from a symmetric quadrature amplitude modulation/phase shift keying (QAM/PSK) constellation is  $\log n$  for large  $n$ , where  $n$  is the number of subcarriers. Therefore, the question is how much reduction beyond  $\log n$  can one asymptotically achieve with coding, and what is the price in terms of the rate loss? In this paper, by optimally choosing the sign of each subcarrier, we prove the existence of  $q$ -ary codes of constant PMEPR for sufficiently large  $n$  and with a rate loss of at most  $\log_q 2$ . We also obtain a Varsharmov–Gilbert-type upper bound on the rate of a code, given its minimum Hamming distance with constant PMEPR, for large  $n$ . Since ours is an existence result, we also study the problem of designing signs for PMEPR reduction. Motivated by a derandomization algorithm suggested by Spencer, we propose a deterministic and efficient algorithm to design signs such that the PMEPR of the resulting codeword is less than  $c \log n$  for any  $n$ , where  $c$  is a constant independent of  $n$ . For symmetric  $q$ -ary constellations, this algorithm constructs a code with rate  $1 - \log_q 2$  and with PMEPR of  $c \log n$  with simple encoding and decoding. Simulation results for our algorithm are presented.

**Index Terms**—Coding, multicarrier signals, peak-to-mean envelope power ratio (PMEPR), orthogonal frequency division multiplexing (OFDM).

## I. INTRODUCTION

RECENTLY, there has been considerable interest in using multicarrier modulation for high-speed communications [2]. As wireline applications, we can mention discrete multi-tone (DMT) in the asymmetric digital subscriber line (ADSL) and the very high rate digital subscriber line (VDSL). Similarly, orthogonal frequency division multiplexing (OFDM) has been proposed for different wireless scenarios such as the wireless local area network (WLAN) and digital video broadcasting (DVB) [2].

In this modulation, information is carried on several narrow-band orthogonal subcarriers, each subcarrier being modulated by a complex constellation like quadrature amplitude modulation (QAM) or phase shift keying (PSK). A major drawback of

using several subcarriers is spurious high-amplitude peaks of the transmitted signal when all of the subcarriers add constructively. To be more specific, considering a multicarrier system with  $n$  subcarriers and each subcarrier being modulated by binary PSK (BPSK) constellation, the worst-case peak-to-mean envelope power ratio (PMEPR) of this system is  $n$  as all the subcarriers add up coherently.

However, several authors noticed that this worst-case PMEPR rarely occurs [3]–[6]. In fact, for large values of  $n$ , it has been recently shown that with probability one, the PMEPR of any randomly chosen modulating vector, carved from any symmetric QAM/PSK constellation, is  $\log n$  asymptotically [6]. Therefore, even though the worst-case PMEPR can be as bad as  $n$ , in the probability sense and in the limit, PMEPR behaves as  $\log n$ .

Several schemes have been proposed to reduce the high PMEPR of multicarrier signals including probabilistic methods (e.g., selective mapping), coding, clipping, and reserved subcarriers [5], [7]–[10]. While existing coding methods give a guarantee on the PMEPR of the system with a large rate hit for large  $n$ , probabilistic methods usually improve the statistical properties of the PMEPR with a little redundancy and using side information [7], [8]. The basic idea behind the probabilistic methods is to lower the probability of occurrence of a peak, and in fact, these methods use the limited redundancy not to eliminate the peaks but only to make them less frequent. Therefore, there is no guarantee on the PMEPR similar to the coding methods.

Recently, Paterson and Tarokh have raised the question of what the tradeoff is between rate and minimum distance of a code with bounded PMEPR [8]. It is also proved that the Varsharmov–Gilbert upper bound remains the same for spherical codes with PMEPR less than  $8 \log n$  for large  $n$ . In [6], based on the asymptotic analysis of PMEPR, it is further shown that the PMEPR of spherical codes and symmetric QAM/PSK constellations is  $\log n$ . However, without contradicting the result of [6], there still might be exponentially many codewords with constant PMEPR, even though the probability of randomly choosing one of them goes to zero, and therefore, they are rare.

In this paper, we start with addressing the achievable PMEPR reduction by choosing an optimum sign for each subcarrier. Based on an elegant result of Spencer on bounded linear forms [11], we prove that by choosing an optimum sign for each subcarrier, we can indeed achieve constant (independent of  $n$ ) bounded PMEPR for sufficiently large  $n$ . Moreover, we find an upper bound for the best constant, and we then use this result and prove the existence of codes carved from a symmetric  $q$ -ary constellation with constant bounded PMEPR and rate greater than  $1 - \log_q 2$ . We also derive a Varsharmov–Gilbert upper

Manuscript received May 21, 2003; revised October 28, 2003. This work was supported in part by the National Science Foundation under Grant CCR-0133818, by the Office of Naval Research under Grant N00014-02-1-0578, and by Caltech's Lee Center for Advanced Networking. The associate editor coordinating the review of this paper and approving it for publication was Prof. Nicholas D. Sidiropoulos.

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: masoud@systems.caltech.edu; hassibi@systems.caltech.edu).

Digital Object Identifier 10.1109/TSP.2004.834343

bound on the rate of a code given its minimum Hamming distance and with constant bounded PMEPR. A scheme with more degrees of freedom is then considered to reduce the PMEPR at the price of further reducing the rate of the code. It is also worth mentioning that this scheme can be interpreted as reducing the PMEPR by expanding the constellation by a factor of two.

We further look into the design of optimum signs to reduce the PMEPR. In order to design the signs, we know that any randomly chosen sign will achieve a PMEPR of  $\log n$  in the probability sense and for large  $n$  [6]. However, randomized algorithms cannot give a guarantee on PMEPR. Therefore, motivated by the derandomization algorithm in [1], we propose a simple algorithm to design signs that guarantees a PMEPR of  $c \log n$  for any  $n$ , where  $c$  is a constant independent of  $n$ . This algorithm in fact constructs a code family with rate loss of  $\log_q 2$  and a PMEPR of less than or equal to  $c \log n$  by using a symmetric constellation with  $q$  alphabets for any  $n$ . Simulation results show significant PMEPR reduction compared with the selective mapping (SLM) method by using Hadamard rows as random signs. Furthermore, our approach allows us to study the peak-to-average power ratio in multiple antenna systems employing linear dispersion (LD) codes [12].

The paper is outlined as follows. Section II introduces our notations and definitions. Section III discusses the peak reduction methods by choosing an optimum sign for each carrier and elaborates the statement of the problem. Furthermore, Section III reviews the mathematical results on bounded linear forms that will be used in the paper. Then, in Section IV, we address the achievable PMEPR reduction by choosing signs where we prove the existence of codes with constant bounded PMEPR for sufficiently large  $n$ . Section V addresses the design of such signs and presents an algorithm to achieve PMEPR of  $c \log n$  for any  $n$ . Section V consequently discusses another application of our method to the peak to average power ratio reduction in multiple antenna systems using LD codes. Finally, Section VI concludes the paper.

## II. DEFINITION

Assuming no pulse shaping, we may represent the complex envelope of a multicarrier signal with  $n$  subcarriers as

$$s_C(t) = \sum_{i=1}^n c_i e^{j2\pi i f_0 t}, \quad 0 \leq t \leq \frac{1}{f_0} \quad (1)$$

where  $f_0$  is the subchannel spacing, and  $C = (c_1, \dots, c_n)$  is the complex modulating vector with entries from a given complex constellation  $\mathcal{Q}$ . The admissible modulating vectors are called codewords, and the ensemble of all possible codewords constitute the code  $\mathcal{C}$ . For mathematical convenience, we define the normalized complex envelope of a multicarrier signal as

$$s_C(\theta) = \sum_{i=1}^n c_i e^{j\theta i}, \quad 0 \leq \theta < 2\pi. \quad (2)$$

Then, the PMEPR of each codeword  $C$  in the code family  $\mathcal{C}$  may be defined as

$$\text{PMEPR}_{\mathcal{C}}(C) = \max_{0 \leq \theta < 2\pi} \frac{|s_C(\theta)|^2}{E\{|C|^2\}}. \quad (3)$$

Similarly,  $\text{PMEPR}_{\mathcal{C}}$  is defined as the maximum of (3) over all codewords in  $\mathcal{C}$ . Clearly, when  $\mathcal{C}$  is a random code such that the  $c_i$ 's are chosen independently from a constellation with average power  $E_{\text{av}}$ , the average power of  $C$  is  $nE_{\text{av}}$ . Throughout the paper, whenever we drop  $\mathcal{C}$  from PMEPR, we mean  $\mathcal{C}$  is the random code with average power  $nE_{\text{av}}$ .

We will use the following notation:  $\mathcal{C}$  and  $C$  represent the code family and codeword, respectively,  $c_i$  denotes the  $i$ th coordinate of the modulating vector  $C$ , and  $\log\{\cdot\}$  is the natural logarithm. We also define the rate of a code  $\mathcal{C}$  chosen from a  $q$ -ary constellation as

$$R = \frac{1}{n} \log_q |\mathcal{C}| \quad (4)$$

where  $|\mathcal{C}|$  is the cardinality of the set  $\mathcal{C}$ . Hamming distance is also defined as the number of coordinates in which two codewords are different, and consequently, the minimum Hamming distance of  $\mathcal{C}$  is its minimum over all pairs of codewords [13].

## III. PEAK REDUCTION SCHEME AND BOUNDED LINEAR FORMS

In this section, for any codeword  $C = (c_1, \dots, c_n)$ , we study designing optimum signs for each subcarrier in order to reduce the PMEPR of  $C$ . We initially motivate and elaborate the statement of the problem, and then, we introduce results on bounded linear forms that will be used in Sections IV–VI to prove the existence of codes with high rate and constant bounded PMEPR.

### A. Peak Reduction by Choosing Optimum Signs

Given the codeword  $C = (c_1, \dots, c_n)$ , we consider the design of optimum sign  $\epsilon_i \in \{+1, -1\}$  for each  $c_i$  in order to minimize the PMEPR of the resulting codeword  $C_{\epsilon} = (\epsilon_1 c_1, \dots, \epsilon_n c_n)$ . Clearly, the worst-case PMEPR of a codeword  $C$  is of the order of  $n$ . We also know that a randomly chosen codeword  $C$  will have a PMEPR of  $\log n$  for large values of  $n$  [6], and therefore, randomly choosing signs should work well for large  $n$ . In fact, randomly choosing signs has shown to be an effective method of PMEPR reduction for moderate values of  $n$  [7]. As an example of similar techniques, in the selective mapping (SLM) method, there are  $M$  statistically independent codewords representing the same information, and the codeword resulting in the lowest PMEPR is selected for transmission; therefore, it needs  $\log_2 M$  bits of side information. This approach was first proposed in [14] for  $M = 2$  and it is generalized in [3] and [7]. For implementation purposes, the  $M$  independent codewords are generated by element-by-element product of the codeword by  $M$  pseudo-random sequences with entries from  $\{+1, -1\}$  or  $\{\pm 1, \pm j\}$ , for instance, Hadamard vectors or  $m$  sequences [15].

This raises the question of how much further reduction in PMEPR we can get by choosing the best sign for each subcarrier. Moreover, is there any deterministic algorithm to design an optimum sign for each subcarrier? Since changing signs does not affect the average power, we can focus on minimizing the peak of  $s_C(\theta)$  over the  $\epsilon_i$ 's. Here is the statement of the problem:

**Problem Statement:** For any given complex vector  $C = (c_1, \dots, c_n)$ , where  $|c_i| \in \mathcal{Q}$  and  $c_i \leq \sqrt{E_{\max}}$ , consider the following minimization problem:

$$\min_{\epsilon} \max_{0 \leq \theta \leq 2\pi} \left| \sum_{i=1}^n \epsilon_i c_i e^{j\theta i} \right| \quad (5)$$

where  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ , and  $\epsilon_i \in \{+1, -1\}$ .

- What is the value of the min-max problem of (5) for any codeword  $C$ , where  $c_i \in \mathcal{Q}$ ?
- How much further improvement can we get by considering more elaborate schemes?
- How can we design the optimum vector  $\epsilon$  efficiently?

The goal of this paper is to address all the above questions, namely, the first two questions on the achievable PMEPR reduction is studied in Section IV for sufficiently large  $n$ , and then in Section V, we propose an algorithm to design the sign vector to guarantee the PMEPR of  $c \log n$  for any  $n$ .

To answer the above problem, we need the following Lemma.

**Lemma 1:** Let  $s_C^R(\theta)$  and  $s_C^I(\theta)$  be the real and imaginary parts of  $s_C(\theta)$  respectively. In addition, let  $\theta_i = (2\pi i/kn)$  for  $i = 1, \dots, kn$ , where  $k > 1$  is such that  $kn$  is an integer. Then

$$\begin{aligned} \max_{0 \leq \theta \leq 2\pi} |s_C(\theta)| \\ \leq \frac{1}{\cos \frac{\pi}{2k}} \sqrt{\max_{1 \leq i \leq kn} |s_C^R(\theta_i)|^2 + \max_{1 \leq i \leq kn} |s_C^I(\theta_i)|^2}. \end{aligned}$$

*Proof:* It is shown by Ehlich and Zeller that a real trigonometric polynomial with  $n$  subcarriers satisfies the following inequality [16]:

$$\max_{0 \leq \theta \leq 2\pi} |s_C^R(\theta)| \leq \frac{1}{\cos \frac{\pi}{2k}} \max_{1 \leq i \leq kn} |s_C^R(\theta_i)| \quad (6)$$

where the  $\theta_i$ 's are  $kn$  uniform samples in  $[0, 2\pi]$ . It is worth noting that various versions of the inequality in (6) have appeared in [17] and [18]. Similarly, the same inequality is valid for the real trigonometric polynomial  $s_C^I(\theta)$ . Now, considering that

$$\max_{0 \leq \theta \leq 2\pi} |s_C(\theta)| \leq \sqrt{\max_{1 \leq i \leq kn} |s_C^R(\theta_i)|^2 + \max_{1 \leq i \leq kn} |s_C^I(\theta_i)|^2} \quad (7)$$

the Lemma follows immediately from (6) and (7). ■

Lemma 1 reformulates the problem in (5) and allows us to consider the optimum  $\epsilon$  to minimize  $2kn$  linear forms constructed by  $s_C^R(\theta_i)$  and  $s_C^I(\theta_i)$  for  $i = 1, \dots, kn$  instead of minimizing the maximum of  $|s_C(\theta)|$  over a continuous variable  $\theta \in [0, 2\pi]$ . Therefore, instead of (5), we may then consider the following minimization problem:

$$\min_{\epsilon} \max_{1 \leq p \leq 2kn} \left| \sum_{i=1}^n \epsilon_i a_{pi} \right| \quad (8)$$

where  $a_{pi}$  is defined as

$$a_{pi} = \begin{cases} \operatorname{Re} \{c_i e^{j\theta_p i}\}, & 1 \leq p \leq kn \\ \operatorname{Im} \{c_i e^{j\theta_p i}\}, & kn + 1 \leq p \leq 2kn \end{cases} \quad (9)$$

where  $\theta_p = (2\pi p/kn)$ . Of course, the optimization problems in (8) and (5) are identical in the limit for large  $k$ . In Section III-B, we briefly review some results on bounded linear forms that will be used to solve the problem in (8).

### B. Bounded Linear Forms

In what follows, we define bounded linear forms and introduce the result of Spencer [11] on the discrepancy of sets with two colorings that can be interpreted as bounding linear forms by using optimum signs.

**Definition 1:** For any  $1 \leq l \leq m$ , linear forms  $L_l$  in  $n$  variables  $(x_1, \dots, x_n)$  are defined as

$$L_l(x_1, \dots, x_n) = \sum_{j=1}^n a_{lj} x_j, \quad 1 \leq l \leq m \quad (10)$$

where all  $a_{lj}$  are real, and  $|a_{lj}| \leq 1$ .

Throughout the paper, we assume  $m \geq n$ , and we define  $\alpha = n/m$ . We quote the following result from [11] to bound the linear forms assuming  $x_j \in \{+1, -1\}$  in Definition 1.

**Theorem 1 (Spencer [11]):** Let  $L_l$  be  $m = \alpha n$  linear forms as defined in Definition 1, where  $\alpha$  is a constant independent of  $n$ . Then, for sufficiently large values of  $n$ , there exist  $\epsilon_1, \dots, \epsilon_n \in \{+1, -1\}$  such that

$$|L_l(\epsilon_1, \dots, \epsilon_n)| \leq \mathcal{K}(\alpha) \sqrt{m} = \frac{\mathcal{K}(\alpha)}{\sqrt{\alpha}} \sqrt{n} \quad (11)$$

where  $\mathcal{K}(\alpha)$  is a constant independent of  $n$  and is bounded by

$$\mathcal{K}(\alpha) \leq 11 \sqrt{\alpha \log(2\alpha^{-1})}. \quad (12)$$

*Proof:* Refer to [11, th. 4 and 7]. ■

Theorem 1 states that there exists a sign vector  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$  that can reduce the maximum of  $m$  linear forms to as much as  $\mathcal{K}(\alpha) \sqrt{m}$  and provides an upper bound for this best constant  $\mathcal{K}(\alpha)$ . In the Appendix, we obtain tighter bounds for  $\mathcal{K}(\alpha)$ ; for instance, the bound for  $\mathcal{K}(\alpha)$  for  $\alpha = 0.5$  in (12) is improved from 9.15 to 4.03. Fig. 1 compares the upper bound of  $(\mathcal{K}(\alpha)/\sqrt{\alpha})$  derived in Appendix A with (12) for different values of  $\alpha$ .

It is also worth mentioning that considering the matrix  $[a_{ij}]$  as a Hadamard matrix, it is clear that there is no sign vector to further reduce the linear forms below  $\sqrt{n}$ . Therefore, an easy lower bound for  $\mathcal{K}(\alpha)/\sqrt{\alpha}$  in Theorem 1 would be one for any  $\alpha$ . Another implication of this is that Hadamard vectors can be considered as good candidates for generating pseudorandom sign vectors for probabilistic methods such as SLM.

**Remark 1:** Note that each codeword  $C$  generates a matrix  $[a_{ij}]$  according to (9). In [11], it is shown that for any  $[a_{ij}]$ , hence any codeword  $C$ , there exist exponentially many vectors  $\epsilon$ ,  $(1+\delta_\alpha)^n$  of them say, such that (11) holds. If one chooses only one of these  $\epsilon$  vectors, then the rate hit is  $\log_q 2$ . However, if we have the choice of choosing different vectors  $\epsilon$  for each codeword, this choice will carry information, and therefore, the rate hit will be  $\log_q 2 - \log_q(1+\delta_\alpha)$ . Characterization of  $\delta_\alpha$  will further reveal the tradeoff between PMEPR and rate. Although having the option of choosing different vector  $\epsilon$  increases the

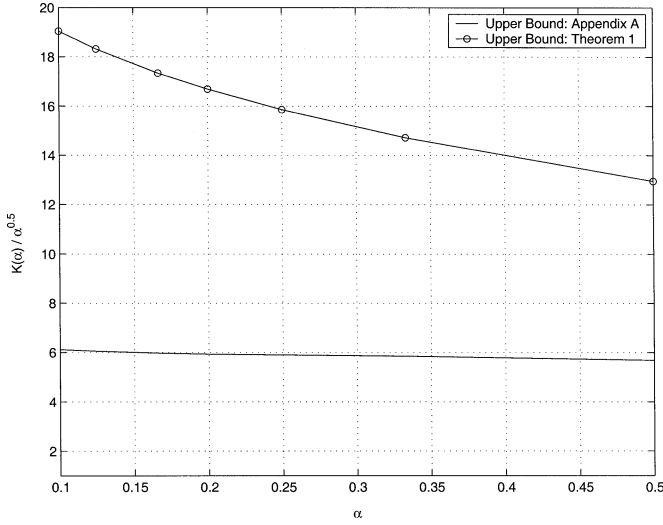


Fig. 1. Upper bounds for  $\mathcal{K}(\alpha)/\sqrt{\alpha}$  for different values of  $\alpha$ .

rate, it greatly complicates the encoding and even more so the decoding, and therefore, we will not consider it further.

#### IV. CODES WITH CONSTANT PMEPR

In this section, based on the scheme presented in Sections III-A and the results in B, we address the problem stated in (5). We study the existence of codes with constant PMEPR for sufficiently large values of  $n$ . We first derive a lower bound on the rate of codes with constant bounded PMEPR, and then, we obtain a Varsharmov–Gilbert bound for the rate and minimum distance of such codes. We further reduce the best achievable constant for PMEPR by choosing the optimum sign for each subcarrier at the price of reducing the rate by using a scheme with more degrees of freedom. The next theorem answers the first question raised in Section III-A.

**Theorem 2:** For any codeword  $C = (c_1, \dots, c_n)$  chosen from a constellation  $\mathcal{Q}$  with maximum and average power  $E_{\max}$  and  $E_{\text{av}}$ , respectively, there exists a vector  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$  with  $\epsilon_i \in \{+1, -1\}$  such that

$$\text{PMEPR}(C_\epsilon) \leq \min_k \left\{ \frac{\sqrt{2k}\mathcal{K}\left(\frac{1}{2k}\right)}{\cos \frac{\pi}{2k}} \right\}^2 \frac{2E_{\max}}{E_{\text{av}}} \quad (13)$$

for sufficiently large  $n$ , where  $C_\epsilon = (\epsilon_1 c_1, \dots, \epsilon_n c_n)$ , and  $k > 1$  is such that  $kn$  is an integer.

*Proof:* To prove the theorem, we use Lemma 1 and then Theorem 1 to bound the linear forms in (8). Therefore, setting  $\alpha = (1/2k) < 1$ , where  $k$  is as in Lemma 1, we can use Theorem 1 to get

$$\max_{0 \leq \theta \leq 2\pi} |s_{C_\epsilon}(\theta)| \leq \frac{\sqrt{2k}\mathcal{K}\left(\frac{1}{2k}\right)}{\cos \frac{\pi}{2k}} \sqrt{2E_{\max}n}. \quad (14)$$

Theorem 2 follows by using the definition of PMEPR. ■

The following Corollary is an immediate consequence of Theorem 2 and states that there exist exponentially many codewords with constant bounded PMEPR for any  $n$  sufficiently large.

**Corollary 1:** For any  $q$ -ary symmetric constellation in which both  $A$  and  $-A$  are in the constellation, there exists a code with

rate  $1 - \log_q 2$  and with constant PMEPR for any  $n$  sufficiently large.

*Proof:* We first consider half of the points in the constellation by choosing one of each two symmetric points, i.e.,  $A$  or  $-A$ . We then use Theorem 2 to state the existence of a choice between any constellation point and its symmetric point such that the PMEPR of the resulting codeword is less than

$$\min_k \left\{ \frac{\sqrt{2k}\mathcal{K}\left(\frac{1}{2k}\right)}{\cos \frac{\pi}{2k}} \right\}^2 \frac{2E_{\max}}{E_{\text{av}}} \quad (15)$$

where  $k$ ,  $E_{\text{av}}$ , and  $E_{\max}$  are as defined in Theorem 2. Therefore, we have at least  $(q/2)^n$  codewords with PMEPR less than (15); thus,  $R \geq 1 - \log_q 2$ . ■

In the following corollary, we obtain a region for the rate and minimum Hamming distance of codes with constant PMEPR.

**Corollary 2:** Let  $\mathcal{Q}_q$  be a symmetric  $q$ -ary constellation as in Corollary 1,  $R > 0$ ,  $q > 2$ , and  $0 \leq \delta \leq (q - 2/q)$ , if we have

$$R \leq (1 - H_{q/2}(\delta))(1 - \log_q 2) \quad (16)$$

where  $H$  is the entropy function. Then asymptotically, there exists a code  $\mathcal{C}$  of length  $n$ , with entries from  $\mathcal{Q}_q$ , rate  $R$ , minimum Hamming distance  $d_{\min} = \lfloor \delta n \rfloor$ , and constant bounded PMEPR.

*Proof:* We first consider one point from every two symmetric points of the constellation, and then, we use the Varsharmov–Gilbert argument for the  $q/2$ -ary constellation to construct a code with rate  $1 - H_{q/2}(\delta)$  and minimum Hamming distance of  $\lfloor \delta n \rfloor$  [13]. Therefore, we can state that if (16) is valid, then there exists a code  $\mathcal{C}'$  of length  $n$ , with entries from the  $q/2$ -ary constellation, rate  $R$ , and minimum Hamming distance  $d_{\min} = \lfloor \delta n \rfloor$ .

Now, we construct the code  $\mathcal{C}$  by modifying the code  $\mathcal{C}'$ . For every codeword  $C = (c_1, \dots, c_n)$  in  $\mathcal{C}'$ , we choose between  $c_i$  and  $-c_i$  in the constellation in order to minimize the PMEPR, and clearly, this does not decrease the minimum Hamming distance of the code  $\mathcal{C}$  from that of  $\mathcal{C}'$ . From Theorem 2, we know that for any codeword in  $\mathcal{C}'$ , there exists such a choice that has PMEPR of less than (15). This completes the proof. ■

We can also consider more sophisticated modifications to a codeword, i.e., choosing between four symmetric points with respect to the imaginary and real axis instead of two, as in Theorem 2. To clarify the idea, let us assume that  $c_i = a_i + jb_i$  and that  $\epsilon_i$ 's and  $\epsilon'_i$ 's are chosen from  $\{+1, -1\}$ , and let

$$s_{C_{\epsilon, \epsilon'}}(\theta) = \sum_{i=1}^n \epsilon_i a_i \cos i\theta - \epsilon_i b_i \sin i\theta + j \sum_{i=1}^n \epsilon'_i a_i \sin i\theta + \epsilon'_i b_i \cos i\theta. \quad (17)$$

In this scheme, when we want to transmit the symbol  $a_i + jb_i$ , we are allowed to use  $\pm a_i$  and  $\pm b_i$  independently in order to reduce the PMEPR. In other words, since this scheme certainly has more degrees of freedom, we can further reduce the PMEPR at the price of reducing the number of codewords to  $(q/4)^n$  for symmetric constellations with respect to the  $x$  and  $y$  axis. The following Corollary addresses this tradeoff between rate and PMEPR, which is analogous to Corollary 1 and 2.

**Corollary 3:** For any  $q$ -ary constellation such that the constellation points are symmetric with respect to both real and imaginary axes, there exists a code with rate at least  $1 - \log_q 4$  and with constant bounded PMEPR for sufficiently large values of  $n$ .

*Proof:* We prove this along the same line as Corollary 1. Since, in this case, we have twice as many degrees of freedom, (14) can be rewritten as

$$\max_{0 \leq \theta \leq 2\pi} |s_{C_{\epsilon, \epsilon'}}(\theta)| \leq \frac{\sqrt{k} \mathcal{K}(\frac{1}{k})}{\cos \frac{\pi}{2k}} \sqrt{2n E_{\max}}. \quad (18)$$

This can be then optimized over  $k$ , defined as in Theorem 2, and therefore, the PMEPR is less than

$$\min_k \left\{ \frac{\sqrt{k} \mathcal{K}(\frac{1}{k})}{\cos \frac{\pi}{2k}} \right\}^2 \frac{2E_{\max}}{E_{av}}. \quad (19)$$

Clearly (19) is less than (15), and this scheme can improve the best achievable constant PMEPR of Corollary 1. On the other hand, since at each time when we want to choose  $a_i + jb_i$ , we use any of the four combinations of  $\pm a_i \pm jb_i$ , and the number of codewords will be at least  $(q/4)^n$ ; therefore, the rate should be greater than  $1 - \log_q 4$ . ■

As shown in Corollary 1, using the optimum signs for each carrier  $\epsilon_i \in \{+1, -1\}$ , we can reduce the PMEPR from  $n$  to a constant of (15) by the rate hit of  $\log_q 2$ . Another interpretation of this is that by expanding the constellation by a factor of 2, we can reduce the PMEPR from  $n$  to a constant for asymptotically large values of  $n$ . By using more degrees of freedom as in Corollary 3, we can further reduce the constant bound for PMEPR from (15) to (19) at the price of the same rate hit, i.e.,  $\log_q 2$ . As Fig. 1 suggests, since the upper bound for  $\mathcal{K}(\alpha)$  is quite close to  $\mathcal{K}(2\alpha)$ , further reduction of the PMEPR by using more degrees of freedom does not seem to be very efficient.

**Remark 2:** Corollary 1 has an interesting implication for the PMEPR distribution. Assuming  $\mathcal{Q}$  to be a symmetric constellation, it is shown in [19] that if the  $c_i$ 's are independently chosen from  $\mathcal{Q}$ , then the OFDM signal tends to a Gaussian process for large  $n$ . Based on this, it is claimed in [19] that the distribution of the PMEPR can be approximated by

$$\Pr\{\text{PMEPR}(C) < \lambda\} \cong e^{-e^{-\lambda} n \sqrt{(\pi/3) \log n}} \quad (20)$$

for large  $n$  and where  $C = (c_1, \dots, c_n)$ . However, (20) cannot be true since it implies that the number of codewords that have a constant PMEPR of  $\lambda$  (independent of  $n$ ) is given by

$$q^{n(1 - (e^{-\lambda} / \log q) \sqrt{(\pi/3) \log n})} \quad (21)$$

which clearly goes to zero as  $n$  tends to infinity. This contradicts Corollary 1. Therefore, even though the OFDM signal is a Gaussian process (any finite number of time samples are jointly Gaussian), this does not say anything about the distribution of the peak since it involves an infinite number of samples.

## V. DESIGN OF SIGNS TO REDUCE THE PMEPR

In this section, we are seeking an answer to the third question raised in Section III, namely, the efficient design of signs

to reduce the PMEPR. As mentioned in Section IV, there exist sign vectors that yield constant bounded PMEPR for sufficiently large values of  $n$ . On the other hand, any random sign vector should have PMEPR of  $\log n$  for large values of  $n$  with probability one, and therefore, random methods should work well in the probability sense. In what follows, we propose a deterministic and efficient algorithm that basically derandomizes the search for the sign vector  $\epsilon$ , and then, we prove that our algorithm guarantees deterministically (*not probabilistically*) that the PMEPR is less than  $c \log n$  for any  $n$  (*not asymptotically*).

We again use Lemma 1 to reformulate the problem in (5), and, instead of designing the vector  $\epsilon$  to minimize the maximum of  $|s_C(\theta)|$  over a continuous variable  $\theta$ , we find the optimum  $\epsilon_i$ 's to minimize  $2kn$  linear forms corresponding to  $s_C^R(\theta_i)$  and  $s_C^I(\theta_i)$  for  $i = 1, \dots, n$  and defined as in (8).

In order to solve (8), we consider a more general setting for our problem. Let us consider the set of equiprobable vectors  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ ,  $\epsilon_i \in \{-1, +1\}$ . Then, for any codeword  $C$ , we define  $A_p^\lambda$  as the event that the  $p$ th linear form defined in (8) is greater than  $\lambda$ . Furthermore, we assume  $\lambda$  is chosen such that  $\sum_{i=1}^{2kn} \Pr\{A_i^\lambda\}$  is less than 1, and therefore, there exists a vector  $\epsilon$  with the above property. We would like to efficiently find the vector  $\epsilon$  such that none of the bad events  $A_p^\lambda$  occur.

This problem has been considered in mathematics and is usually referred to as the derandomization of random algorithms [1], [20]. In this approach, we assume that we can compute the conditional probability  $\Pr\{A_p^\lambda | \epsilon_1, \dots, \epsilon_j\}$ , and we find the  $\epsilon_i$ 's sequentially. At the  $j$ th step, given the optimally chosen signs  $\epsilon_1^*, \dots, \epsilon_{j-1}^*$ , we choose  $\epsilon_j^* \in \{+1, -1\}$  such that

$$\begin{aligned} & \sum_{i=1}^{2kn} \Pr\{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j^*\} \\ &= \min \left\{ \sum_{i=1}^{2kn} \Pr\{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = +1\} \right. \\ & \quad \left. \sum_{i=1}^{2kn} \Pr\{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = -1\} \right\}. \quad (22) \end{aligned}$$

Due to the above recursive minimization and assuming that  $\epsilon_1^*, \dots, \epsilon_{j-1}^*$  are determined, we can write the inequality (23), shown at the bottom of the next page. Finally, since  $\sum_{i=1}^{2kn} \Pr\{A_i^\lambda\} < 1$ , and after determining each  $\epsilon_j$  sequentially according to (22), we will end up with

$$\sum_{i=1}^{2kn} \Pr\{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_n^*\} < 1. \quad (24)$$

Since there is no randomness in the conditional events of (24) when all the  $\epsilon_i^*$ 's are determined, each  $\Pr\{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_n^*\}$  is either one or zero. Therefore, (24) implies that all of the probabilities are zero, and consequently, the resulting vector  $\epsilon^*$  guarantees that none of the events  $A_i^\lambda$  will occur.

The difficulty here is now in the efficient computation of the conditional probabilities. Instead of using the exact conditional

<sup>1</sup>We use the superscript star to denote that the  $\epsilon_j$ 's are optimally chosen not arbitrary.

probability functions, we can use upper bounds for conditional probabilities defined as

$$\Pr \{A_i^\lambda | \epsilon_1, \dots, \epsilon_j\} \leq F_i^\lambda(\epsilon_1, \dots, \epsilon_j) \quad (25)$$

if the upper bounds satisfy the following conditions:

$$\begin{aligned} i) \quad & \sum_{i=1}^{2kn} F_i^\lambda < 1 \\ ii) \quad & F_i^\lambda(\epsilon_1, \dots, \epsilon_j) \geq \min_{\epsilon_j \in \{+1, -1\}} F_i^\lambda(\epsilon_1, \dots, \epsilon_{j-1}, \epsilon_j). \end{aligned} \quad (26)$$

Obviously, by the same reasoning used for the original algorithm, we can use the upper bound to find the vector  $\epsilon^*$  such that none of the events  $A_i^\lambda$  occur. Fortunately, as will be proved in Theorem 3, Chernoff's bound does the work for us

$$\begin{aligned} & \Pr \left\{ \left| \sum_{i=1}^n a_{pi} \epsilon_i \right| > \lambda | \epsilon_1, \dots, \epsilon_j \right\} \\ & \leq 2e^{-\gamma\lambda} \cosh \left\{ \gamma \sum_{r=1}^j \epsilon_r a_{pr} \right\} \prod_{r=j+1}^n \cosh \gamma a_{pr} \\ & \stackrel{\text{def.}}{=} F_p^\lambda(\epsilon_1, \dots, \epsilon_j) \end{aligned} \quad (27)$$

for any  $\gamma > 0$  and  $1 \leq p \leq 2kn$ . We will show in the proof of Theorem 3 that  $F_p^\lambda(\epsilon_1, \dots, \epsilon_j)$ , as defined in (27), satisfies both conditions in (26). Now, we return to our problem and present the following algorithm.

*Algorithm 1:* For any codeword  $C = (c_1, \dots, c_n)$ , let  $a_{pi}$  be as in (8),  $k$  be as in Lemma 1, and  $|c_i| \leq \sqrt{E_{\max}}$ . Then,  $\epsilon_1 = 1$ , and  $\epsilon_j$ 's are recursively determined as

$$\begin{aligned} \epsilon_j = -\text{sign} \left\{ \sum_{p=1}^{2kn} \sinh \left\{ \gamma^* \sum_{r=1}^{j-1} \epsilon_r a_{pr} \right\} \sinh(\gamma^* a_{pj}) \right. \\ \left. \times \prod_{r=j+1}^n \cosh \{\gamma^* a_{pr}\} \right\} \end{aligned} \quad (28)$$

for  $j = 2, \dots, n$ , where  $\gamma^* = \sqrt{(2 \log 4kn / n E_{\max})}$ . ■

The following Theorem gives the worst-case guarantee on the PMEPR of the codeword  $C_\epsilon = (\epsilon_1 c_1, \dots, \epsilon_n c_n)$ .

*Theorem 3:* Let  $C = (c_1, \dots, c_n)$  be a given codeword, where  $|c_i| \leq \sqrt{E_{\max}}$ , and  $E_{\text{av}} = E\{|c_i|^2\}$ . In addition, let  $C_\epsilon = (\epsilon_1 c_1, \dots, \epsilon_n c_n)$ , where  $\epsilon_i \in \{+1, -1\}$  is determined according to Algorithm 1. Then, the PMEPR of the resulting codeword  $C_\epsilon$  will be less than  $(4E_{\max} / \cos^2(\pi/2k) E_{\text{av}}) \log 4kn$ , where  $k$  is as in Lemma 1.

*Proof:* The proof relies on the derandomization method illustrated before and uses the Chernoff bound to evaluate the conditional probability distributions. As a first step, we derive the upper bound for the conditional probability in (25) as

$$\begin{aligned} & \Pr \left\{ \left| \sum_{i=1}^n a_{pi} \epsilon_i \right| > \lambda | \epsilon_1, \dots, \epsilon_j \right\} \\ & = \Pr \left\{ \sum_{i=1}^n a_{pi} \epsilon_i > \lambda | \epsilon_1, \dots, \epsilon_j \right\} \\ & \quad + \Pr \left\{ \sum_{i=1}^n a_{pi} \epsilon_i < -\lambda | \epsilon_1, \dots, \epsilon_j \right\} \\ & = \Pr \left\{ \sum_{i=j+1}^n a_{pi} \epsilon_i > \lambda - \sum_{i=1}^j a_{pi} \epsilon_i | \epsilon_1, \dots, \epsilon_j \right\} \\ & \quad + \Pr \left\{ -\sum_{i=j+1}^n a_{pi} \epsilon_i > \lambda + \sum_{i=1}^j a_{pi} \epsilon_i | \epsilon_1, \dots, \epsilon_j \right\} \\ & \leq e^{\gamma \sum_{i=1}^j a_{pi} \epsilon_i} \times e^{-\gamma\lambda} E \left\{ e^{\gamma \sum_{i=j+1}^n a_{pi} \epsilon_i} \right\} \\ & \quad + e^{-\gamma \sum_{i=1}^j a_{pi} \epsilon_i} \times e^{-\gamma\lambda} E \left\{ e^{-\gamma \sum_{i=j+1}^n a_{pi} \epsilon_i} \right\} \\ & = 2e^{-\gamma\lambda} \cosh \left\{ \gamma \sum_{r=1}^j \epsilon_r a_{pr} \right\} \prod_{r=j+1}^n \cosh \gamma a_{pr} \\ & = F_p^\lambda(\epsilon_1, \dots, \epsilon_j) \end{aligned} \quad (29)$$

where we used Chernoff's bound and the fact that  $\epsilon_i \in \{+1, -1\}$  is equiprobable and where  $\gamma$  is the Chernoff's bound coefficient to be optimized. We then show that the upper bound satisfies the conditions in (26). Using  $\cosh(a+b) + \cosh(a-b) = 2 \cosh a \cosh b$  and using the definition of  $F_p^\lambda$ , (29) can be written as (30), shown at the

$$\begin{aligned} \sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*\} &= \frac{\sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = 1\} + \sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = -1\}}{2} \\ &\geq \min \left\{ \sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = 1\}, \sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = -1\} \right\} \\ &= \sum_{i=1}^{2kn} \Pr \{A_i^\lambda | \epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j^*\} \end{aligned} \quad (23)$$

bottom of the page, which ensures the second condition in (26) is satisfied. To verify the first condition, we use the following inequalities:

$$\begin{aligned} \sum_{i=1}^{2kn} F_i^\lambda &= \sum_{i=1}^{2kn} 2e^{-\gamma\lambda} \prod_{r=1}^n \cosh \gamma a_{pr} \\ &< \sum_{i=1}^{2kn} 2e^{-\gamma\lambda} \prod_{r=1}^n e^{(\gamma^2 a_{pr}^2/2)} \\ &\leq \sum_{i=1}^{2kn} 2e^{-\gamma\lambda} \times e^{(\gamma^2/2)nE_{\max}} \\ &\leq 4kne^{-\gamma\lambda + (\gamma^2/2)nE_{\max}} \end{aligned}$$

where we used  $\cosh x < e^{x^2/2}$  for  $x \neq 0$ . Now, we can optimize over  $\gamma$  to get  $\gamma^* = \lambda/nE_{\max}$  and setting  $\lambda = \sqrt{2nE_{\max}} \log 4kn$ , (31) can be written as

$$\begin{aligned} \sum_{i=1}^{2kn} \Pr \left\{ \left| \sum_{i=1}^n a_{pi} \epsilon_i \right| > \lambda \right\} \\ \leq \sum_{i=1}^{2kn} F_i^\lambda < 4kne^{-(\lambda^2/2nE_{\max})} = 1. \quad (31) \end{aligned}$$

Therefore, the first condition will be also satisfied. Based on (22), we can compute  $\epsilon_j^*$  in each step given  $\epsilon_1^*, \dots, \epsilon_{j-1}^*$  as

$$\epsilon_j^* = -\text{sign} \left\{ \sum_{i=1}^{kn} F_i^\lambda (\epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = +1) - \sum_{i=1}^{kn} F_i^\lambda (\epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j = -1) \right\}. \quad (32)$$

Substituting the definition of  $F_i^\lambda (\epsilon_1^*, \dots, \epsilon_{j-1}^*, \epsilon_j)$  in (32),  $\epsilon_j^*$  can be written as in (28). Lemma 1 then completes the proof by setting the value of  $\lambda$  and relating PMEPR to the maximum of  $2kn$  linear forms, and it can be easily shown that the PMEPR of the resulting codeword is less than  $(4E_{\max}/\cos^2(\pi/2k)E_{\text{av}}) \log 4kn$ , where  $k$  is as in Lemma 1. ■

In order to get better insight to the above result, we define the rate of a  $q$ -ary code family  $\mathcal{C}$  as in (4). In fact, Theorem 3 implies that by using an optimum sign for each subcarrier,

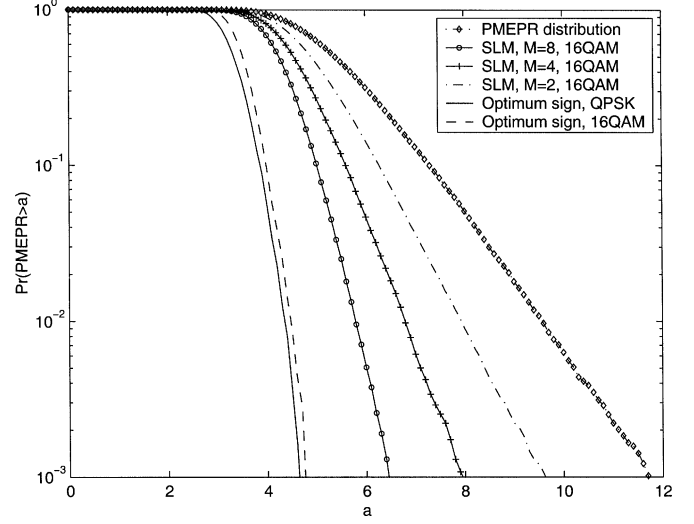


Fig. 2. PMEPR distributions for  $n = 64$  and for different schemes: SLM with  $M = 2, 4$ , and  $8$  and using signs based on Algorithm 1 for each subcarrier.

we can construct a code with rate  $1 - \log_q 2$  and PMEPR of  $c \log n$  for any  $n$ . The rate and PMEPR of this code is much higher than those of the previous codes proposed in [8] (and references therein) whose PMEPR is  $O(\log^2 n)$  and whose rate is approaching to zero as  $n$  increases. It is also worth mentioning that finding optimum signs in the transmitter side can be done very efficiently, and the decoding is very simple since the decoder simply ignores the sign of each subcarrier. On the other hand, this scheme can be interpreted as a scheme to reduce the PMEPR by expanding the constellation. For instance, by expanding the number of constellation points by a factor of two and making it symmetric, the resulting PMEPR can be reduced from  $n$  to the order of  $\log n$  for any  $n$ .

#### A. Simulation Results

In this section, we present simulation results for different constellations, including quadrature phase shift keying and 16QAM and for  $n = 64$  and  $n = 128$ . Fig. 2 shows the actual complementary cumulative distribution function (CCDF) of PMEPR ( $\Pr\{\text{PMEPR} > a\}$ ) and compares it with the PMEPR distributions after using SLM method by Hadamard vectors as pseudorandom sequences for  $M = 2, 4$ , and  $8$ , i.e., the transmitter

$$\begin{aligned} F_p^\lambda(\epsilon_1, \dots, \epsilon_{j-1}) &= 2e^{-\gamma\lambda} \cosh \left\{ \gamma \sum_{r=1}^{j-1} \epsilon_r a_{pr} \right\} \cosh \gamma a_{pj} \prod_{r=j+1}^n \cosh \gamma a_{pr} \\ &= e^{-\gamma\lambda} \prod_{r=j+1}^n \cosh \gamma a_{pr} \left( \cosh \left\{ \gamma \sum_{r=1}^{j-1} \epsilon_r a_{pr} + \gamma a_{pj} \right\} + \cosh \left\{ \gamma \sum_{r=1}^{j-1} \epsilon_r a_{pr} - \gamma a_{pj} \right\} \right) \\ &= \frac{F_p^\lambda(\epsilon_1, \dots, \epsilon_{j-1}, \epsilon_j = +1) + F_p^\lambda(\epsilon_1, \dots, \epsilon_{j-1}, \epsilon_j = -1)}{2} \\ &\geq \min_{\epsilon_j \in \{+1, -1\}} F_i^\lambda(\epsilon_1, \dots, \epsilon_{j-1}, \epsilon_j). \end{aligned} \quad (30)$$

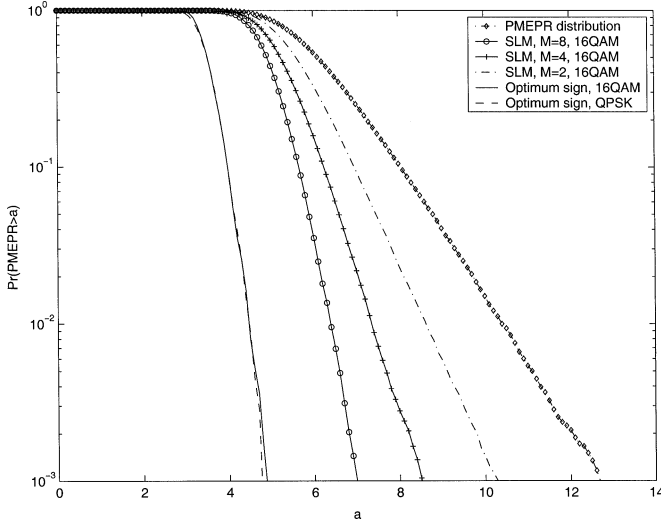


Fig. 3. PMEPR distributions for  $n = 128$  and for different schemes: SLM with  $M = 2, 4$ , and  $8$  and using signs based on Algorithm 1 for each subcarrier.

sends the best codeword out of  $M$  in terms of PMEPR. Fig. 2 also shows the CCDF of PMEPR after using the signs derived by Algorithm 1. Clearly, the distribution function improves significantly. For instance, the probability of having PMEPR of 4.6 is almost one; however, by using signs based on Algorithm 1, this probability will go down to  $10^{-3}$ .

Fig. 3 compares similar schemes when the number of subcarriers is 128. Interestingly, the gain in PMEPR reduction here is much more. Furthermore, the PMEPR distribution after using the designed signs for  $n = 64$  and  $n = 128$  are very close. It is also worth noting that the PMEPR drop is much more abrupt after optimizing the signs. For example, the probability of having PMEPR greater than 4 is almost one; however, the probability of having PMEPR greater than 4.8 goes down to  $10^{-3}$ .

## VI. OTHER APPLICATIONS

Since the results in this paper are more general than just for multicarrier signals, we can address the PMEPR problem for non-Harmonic multicarrier signals as a straightforward generalization. We can also consider the joint PMEPR and intercarrier interference (ICI) reduction in OFDM systems that has been recently addressed in [21]. In Section VI-A, we consider the peak to average power ratio (PAPR) in multiple antenna systems employing LD codes [12].

### A. PAPR in Multiple Antenna Systems Using LD Codes

Another issue of interest is the peak value analysis of the transmitted signal in multiple antenna systems. We consider LD codes for two reasons: First, LD codes are very general and includes many proposed codes, and second, LD codes are linear as a function of the information symbols, and this allows us to use our approach to address this problem.

We consider the following systems: Assume  $M$  and  $N$  are the number of transmit and receive antennas and that  $\alpha_i$ 's and

$\beta_i$ 's are chosen from a constellation. Then, using LD codes, the transmit matrix  $S = [s_{ij}]_{T \times M}$  over  $T$  time slots is defined as

$$S = \sum_{q=1}^Q \alpha_q A_q + j \beta_q B_q \quad (33)$$

where  $A_q$ 's are constant  $T \times M$  matrices to be optimized to achieve capacity and where  $Q \leq \min(M, N)T$ . Since each entry of the matrix  $S$  is the linear combination of  $2Q$  independent information symbols,  $|s_{ij}|$ , in the worst case, can be in the order of  $2Q$  as all of the terms add up coherently. To mathematically quantify this effect, we define the peak-to-average power ratio (PAPR) of the multiple antenna system as

$$\text{PAPR}(C) = \frac{\max_{i,j} |s_{ij}|^2}{E \left\{ \sum_{i,j} |s_{ij}|^2 \right\}} = \frac{\max_{i,j} |s_{ij}|^2}{MT} \quad (34)$$

since the average transmit power is  $MT$ , i.e.,  $E\{tr(SS^*)\} = E \left\{ \sum_{i,j} |s_{ij}|^2 \right\} = MT$ , and where  $C = (\alpha_1, \dots, \alpha_Q, \beta_1, \dots, \beta_Q)$  is the information vector that is mapped to the transmit matrix  $S$ , as in (33). Similarly, the maximum of  $\text{PAPR}(C)$  over all the admissible vectors of  $C$  is the PAPR of the multiple antenna system.

To get a better insight on how severe this problem can be for multiple antenna systems, let the  $\alpha_i$ 's and the  $\beta_i$ 's be chosen from a BPSK constellation, and therefore, (33) can be written as

$$S = \sum_{i=1}^Q \alpha_q A_q^R - \beta_q B_q^I + j \sum_{i=1}^Q \alpha_q A_q^I + \beta_q B_q^R \quad (35)$$

where  $A^R$  and  $A^I$  correspond to the real and imaginary parts of the matrix  $A$ . Furthermore, we assume

$$\sum_{j=1}^Q (a_{ij}^R)^2 + (b_{ij}^I)^2 \leq Q \text{ and } \sum_{j=1}^Q (a_{ij}^I)^2 + (b_{ij}^R)^2 \leq Q. \quad (36)$$

Clearly, if (36) is not valid, we can simply scale  $a_{ij}^R$ 's and  $b_{ij}^I$ 's to satisfy (36) without affecting the PAPR, as defined in (34). This assumption simplifies our derivations and makes the result more clear. Clearly, the worst-case analysis can give us the PAPR of  $O(Q)$  as all the terms add up constructively. However, the following theorem shows that encountering PAPR of greater than  $O(\log Q)$  is highly unlikely as  $Q$  is getting large, and  $N/M$  is a constant.

**Lemma 2:** Under the assumption of (36) and when the  $\alpha_i$ 's and  $\beta_i$ 's are chosen from a BPSK constellation

$$\Pr \left\{ \text{PAPR}(C) > \frac{4Q}{MT} \log MT + \frac{4Q}{MT} \log \log MT \right\} \leq \frac{1}{\log 2MT}. \quad (37)$$



*Proof:* Since  $\alpha_i$ 's and  $\beta_i$ 's are independent has uniform distribution over  $\{+1, -1\}$ , we can use (35) to write

$$\begin{aligned} \Pr\{|s_{ij}^R| > \lambda\} &\leq 2\Pr\{s_{ij}^R > \lambda\} \leq 2\frac{E\{e^{ts_{ij}^R}\}}{e^{t\lambda}} \\ &= 2e^{-t\lambda} E\left\{e^{t\sum_{i=1}^Q \alpha_i a_{ij}^R - \beta_i b_{ij}^I}\right\} \end{aligned} \quad (38)$$

where we used the union bound and Chernoff's bound for the first and second inequality. We can now further use the distribution of  $\alpha_i$ 's and  $\beta_i$ 's to bound the characteristic function as

$$\begin{aligned} E\{e^{ts_{ij}^R}\} &= \prod_{i=1}^Q \{\cosh ta_{ij}^R \times \cosh tb_{ij}^I\} \\ &\leq e^{(t^2/2)\sum_{i=1}^Q (a_{ij}^R)^2 + (b_{ij}^I)^2} \leq e^{t^2 Q/2} \end{aligned} \quad (39)$$

where we used the inequality  $\cosh \alpha \leq e^{\alpha^2/2}$  and (36). Therefore, (38) can be simplified to

$$\Pr\{|s_{ij}^R| > \lambda\} \leq 2e^{t^2 Q/2 - t\lambda} \quad (40)$$

for any  $t \geq 0$ , and similarly, the same inequality holds for  $s_{ij}^I$ . We can now optimize over  $t$  to set  $t^* = (\lambda/Q)$ , and use the following inequalities to get

$$\begin{aligned} \Pr\left\{\max_{i,j}|s_{ij}| > \sqrt{2}\lambda\right\} &\leq MT\Pr\{|s_{ij}| > \sqrt{2}\lambda\} \\ &\leq MT\Pr\{|s_{ij}^R| > \lambda\} \\ &\quad + MT\Pr\{|s_{ij}^I| > \lambda\} \\ &\leq 2MTe^{-\lambda^2/2Q}. \end{aligned}$$

The Lemma follows by letting

$$\lambda = \sqrt{2Q \log MT + 2Q \log \log MT}. \quad \blacksquare$$

Lemma 2 states that even though the worst-case PAPR can be of the order of  $Q$ , it is highly unlikely to encounter peaks greater than  $O(\log Q)$  since  $Q = \min\{M, N\} \times T$ .

Now, since the problem here again can be reduced to bounding linear forms, we can follow along the same line as Section V and show that there exist codes with constant PAPR by just choosing optimum signs for each  $\alpha$  and  $\beta$ . In this case, we have a sign vector with  $2Q$  elements and  $2MT$  linear forms. We can also use the algorithm in [22] to find the vector  $\epsilon$  by simply using  $2MT$  linear forms corresponding to the real and imaginary parts of the entries of  $S$  and  $Q$  signs to be chosen to reduce  $\max_{i,j}|s_{ij}|$ .

## VII. CONCLUSION

We proved the existence of  $q$ -ary codes with rate greater than  $1 - \log_q 2$  with constant PMEPR when the number of subcarriers  $n$  is large. In fact, we can achieve this region by using the optimum sign for each subcarrier to reduce PMEPR. We also obtained a Varsharmov–Gilbert upper bound on the rate of a code given its minimum Hamming distance with constant PMEPR. We then considered a scheme to choose between four constel-

lation points (rather than two) to further reduce PMEPR at the price of reducing the rate. In the second part of the paper, we considered the design of signs to reduce the PMEPR. We proposed a deterministic algorithm that computes the optimum sign vector efficiently and guarantees a PMEPR of  $c \log n$  for any  $n$ . This scheme allows us to reduce the PMEPR at the price of expanding the constellation. Simulation results show a large improvement in the PMEPR by expanding the constellation by a factor of 2. Finally, the framework here was used to study the PAPR of multiple antenna systems employing LD codes.

## APPENDIX A

In this Appendix, we obtain a tighter values for  $\mathcal{K}(\alpha)$  than the bound in Theorem 1. We first quote the following theorem from [11], which will be used to find better bounds.

*Theorem A.1 (Spencer [11]-Theorem 10):* Let  $L_t$  be  $m$  linear forms, as defined in Definition 1. Let  $t$ , which is an infinite positive sequence  $\gamma_1, \gamma_2, \dots, \beta$ , and  $p$ , be given, satisfying

$$\begin{aligned} 1) \quad &\beta = \frac{1}{\alpha} \sum_{s=1}^{\infty} H(2\gamma_s Q(t(2s-1))) + 2\gamma_s Q(t(2s-1)) \\ 2) \quad &\sum_{s=1}^{\infty} \gamma_s^{-1} = 1 - c < 1 \\ 3) \quad &H\left(\frac{1}{2} - p\right) < 1 - \beta \end{aligned} \quad (A.1)$$

then there exist  $\epsilon_1, \dots, \epsilon_n \in \{-1, 0, +1\}$  such that

$$\begin{aligned} 1) \quad &|\{i : \epsilon_i = 0\}| \leq 2pn \\ 2) \quad &|L_l(\epsilon_1, \dots, \epsilon_n)| \leq t\sqrt{\alpha}\sqrt{m} \quad 1 \leq l \leq m. \end{aligned} \quad (A.2)$$

It is also worth noting that in the proof of the above theorem, sufficiently large  $n$  means that

$$n \geq \frac{-\log c}{1 - \beta - H\left(\frac{1}{2} - p\right)}$$

where  $\beta$ ,  $c$ , and  $p$  are as defined in Theorem A.1 [11]. Now, we can state the following Lemma.

*Lemma 3:* Let  $\mathcal{K}(\alpha)$  be defined as in Theorem 1. Then, for any  $t > 3$

$$\mathcal{K}(\alpha) \leq t\sqrt{\alpha} + \mathcal{K}\left(\sqrt{-3.05\alpha Q(t) \log_2 0.39Q(t)}\right) \quad (A.3)$$

where  $Q(t) = (1/\sqrt{2\pi}) \int_t^{\infty} e^{-x^2/2} dx$ .

*Proof:* As Theorem 1 states, there exists a vector  $\epsilon$  with entries chosen from  $\{+1, -1\}$  such that all  $m$  linear forms are bounded by  $\mathcal{K}(\alpha)\sqrt{m}$  for sufficiently large values of  $n$ . Now, to bound the linear forms, we first use Theorem A.1 and then use Theorem 1 to assign a sign to the remaining coefficients that were assigned zero by Theorem A.1. Therefore, using the definition of  $\mathcal{K}$  and triangle inequality, we get

$$\mathcal{K}(\alpha) \leq t\sqrt{\alpha} + \mathcal{K}(2p\alpha) \quad (A.4)$$

for sufficiently large  $n$ . From now on, we use Theorem A.1 to compute  $p$  in (A.4) as a function of  $t$  and  $\alpha$ . As a first step, we use the inequality  $H(1/2 - p) \leq 1 - 2.88p^2$  for  $p < 0.5$  to choose  $p$  as a function of  $\beta$  as

$$H\left(\frac{1}{2} - p\right) \leq 1 - 2.88p^2 = 1 - \beta^{up} \leq 1 - \beta \quad (A.5)$$

where  $\beta^{up}$  is an upper bound for  $\beta$ . Therefore, (A.5) guarantees the third condition, and we only need to find an upper bound for  $\beta$  and simplify the first condition in (A.1). In Appendix B, it is further shown that the first and second conditions of (A.1) imply that

$$\alpha\beta \leq -2.2Q(t) \log_2 0.39Q(t) = \alpha\beta^{up} \quad t \geq 3. \quad (\text{A.6})$$

Inserting the upper bound of  $\beta$  in (A.5), we deduce that  $p = \sqrt{(0.763/\alpha)Q(t) \log_2 0.39Q(t)}$ . Lemma follows by substituting  $p$  in (A.4). ■

We can now numerically compute each  $\mathcal{K}(\alpha)$  by using the recursion of Lemma 3. For example, for  $\alpha = 0.5$ , letting  $t = 5$ , we get

$$\mathcal{K}(0.5) \leq 5\sqrt{0.5} + \mathcal{K}(0.0032)$$

again using the Lemma by  $t = 5$

$$\mathcal{K}(0.5) \leq 5\sqrt{0.5} + 6\sqrt{0.0032} + \mathcal{K}(1.75 \times 10^{-5}). \quad (\text{A.7})$$

For the last term, we can use the bound in Theorem 1 (i.e.,  $\mathcal{K}(\alpha) \leq 11\sqrt{\alpha \log 2\alpha^{-1}}$ ), and hence, (A.7) implies that  $\mathcal{K}(0.5) \leq 4.03$ , which is much better than 9.15, as suggested by Theorem 1.

## APPENDIX B

In this Appendix, for  $t > 3$ , we simplify the first condition of (A.1) and reveal the relationship between  $\beta$ ,  $\alpha$ , and  $t$ . The idea is to show that for  $t > 3$ , the dominant term is the first term with  $s = 1$  in (A.1). As a first step, we assume  $\gamma_1 = 1.1$  and  $\gamma_i = 20^{i-1}$  for  $i \geq 2$  to satisfy the second condition in (A.1). Therefore, in the derivation of Lemma 3, we will not optimize over the values of  $\gamma_i$ 's. Now, inserting the value of  $\gamma_i$ 's in the first inequality of (A.1), we may write

$$\begin{aligned} \alpha\beta &\leq H(2.2Q(t)) + 2.2Q(t) \\ &+ \sum_{s=2}^{\infty} \left\{ -2Q(t(2s-1))20^{s-1} \right. \\ &\quad \times \log_2 \{ 2Q(t(2s-1))20^{s-1} \} \\ &\quad \left. + 4Q(t(2s-1))20^{s-1} \right\} \end{aligned} \quad (\text{B.1})$$

where we used the fact that  $H(\epsilon) \leq -\epsilon \log_2 \epsilon + \epsilon \log_2 e$ , which follows from  $\log_2(1-x) \leq x \log_2 e$ . Now, we can further simplify (B.1) by using  $e^{-t^2} \leq Q(t) \leq 1/2e^{-t^2/2}$  for  $t \geq 3$  to obtain

$$\begin{aligned} \alpha\beta &\leq H(2.2Q(t)) + 2.2Q(t) \\ &+ \sum_{s=2}^{\infty} \left\{ -20^{s-1}e^{-(t^2/2)(2s-1)^2} \right. \\ &\quad \times \log_2 \left\{ 2e^{-t^2(2s-1)^2} 20^{s-1} \right\} \\ &\quad \left. + 2.43e^{-(t^2/2)(2s-1)^2} 20^{s-1} \right\} \\ &\leq H(2.2Q(t)) + 2.2Q(t) \\ &+ \sum_{s=2}^{\infty} \left\{ 20^{s-1}e^{-t^2(2s-1)} \left( -\left(s - \frac{1}{2}\right) \log_2 20 \right. \right. \\ &\quad \left. \left. + t^2(2s-1)^2 \right) + 1.43e^{-t^2(2s-1)} 20^{s-1} \right\} \end{aligned} \quad (\text{B.2})$$

where we used  $s - 1/2 < s - 1$  and  $2s - 1 > 2$  for  $s \geq 2$ . The above summation can be easily worked out by defining

$$f(t^2) = \sum_{s=2}^{\infty} 20^{s-1}e^{-t^2(2s-1)} = \frac{20e^{-3t^2}}{1 - 20e^{-2t^2}}. \quad (\text{B.3})$$

Therefore, (B.2) can be written as

$$\begin{aligned} \alpha\beta &\leq H(2.2Q(t)) + 2.2Q(t) + 1.43f(t^2) \\ &+ \frac{1}{2} \log 20 \left( -\frac{df(a)}{da} \Big|_{a=t^2} \right) + t^2 \left( \frac{d^2f(a)}{da^2} \Big|_{a=t^2} \right) \\ &\leq H(2.2Q(t)) + 2.2Q(t) + 28.6e^{-3t^2} \\ &\quad + 60 \log_2 \sqrt{20}e^{-3t^2} + 180t^2e^{-3t^2} \\ &\leq H(2.2Q(t)) + 2.3Q(t) \end{aligned} \quad (\text{B.4})$$

where the second inequality is valid for  $t \geq 3$  by differentiating (B.3) and letting  $t = 3$  in the denominator. For the third inequality, we can simply observe that the last three terms are less than  $e^{-2t^2}$ , and since  $0.1Q(t) > e^{-2t^2}$  for  $t \geq 3$ , the third inequality follows. We can further simplify (B.4) by using  $H(\epsilon) \leq -\epsilon \log \epsilon + \epsilon \log_2 e$  to obtain

$$\begin{aligned} \alpha\beta &\leq H(2.2Q(t)) + 2.3Q(t) \\ &\leq -2.2Q(t) \log_2 \{ 2.2Q(t) \} + 5.43Q(t) \\ &= -2.2Q(t) \log_2 \{ 0.39Q(t) \} \end{aligned} \quad (\text{B.5})$$

which yields (A.6).

## REFERENCES

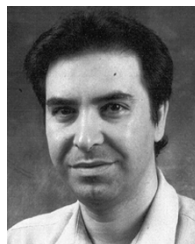
- [1] J. Spencer, "Ten lectures on the probabilistic method," in *Proc. SIAM CBMS-NSF Regional Conf. Series Applied Math.*, 1994.
- [2] A. R. Bahai and B. R. Saltzberg, *Multicarrier Digital Communications: Theory and Applications of OFDM*. New York: Plenum, 1999.
- [3] R. W. Bauml, R. F. H. Fischer, and J. B. Huber, "Reducing the peak to average power ratio of multicarrier modulation by selected mapping," *Electron. Lett.*, vol. 32, no. 22, Oct. 1996.
- [4] H. Ochiai and H. Imai, "On the distribution of the peak to average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, pp. 282-289, Feb. 2001.
- [5] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of OFDM," in *Proc. IEEE Veh. Technol. Conf.*, May 1997, pp. 1634-1638.
- [6] M. Sharif and B. Hassibi, "On multicarrier signals where the PMEPR of a random codeword is  $\log n$ ," *IEEE Trans. Inform. Theory*, vol. 50, pp. 895-903, May 2004.
- [7] S. H. Muller and J. B. Huber, "A comparison of peak power reduction schemes for OFDM," in *Proc. IEEE Global Commun. Conf.*, 1997, pp. 1-5.
- [8] K. G. Paterson and V. Tarokh, "On the existence and construction of good codes with low peak to average power ratios," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1974-1986, Sept. 2000.
- [9] J. A. Davis and J. Jedwab, "Peak to mean power control in OFDM, golay complementary sequences, and reed-muller codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2397-2417, Nov. 1999.
- [10] J. Tellado and J. M. Cioffi, "Efficient algorithms for reducing PAR in multicarrier systems," in *Proc. IEEE Inter. Symp. Inform.*, Aug. 1998, pp. 191-191.
- [11] J. Spencer, "Six standard deviations suffice," *Trans. Amer. Math. Soc.*, vol. 289, no. 2, pp. 679-706, June 1985.
- [12] B. Hassibi and B. Hochwald, "High rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804-1824, July 2002.
- [13] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [14] D. Mestdag and P. Spruyt, "A method to reduce the probability of clipping in DMT based transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 1234-1238, Oct. 1996.
- [15] R. Van Nee, *OFDM for Wireless Multimedia Communication*. Boston, MA: Artech House, 2000.

- [16] H. Ehlich and K. Zeller, "Schwankung von polynomen zwischen gitterpunkten," *Math. Zeitschr.*, vol. 86, pp. 41–44, 1964.
- [17] G. Wunder and H. Boche, "Peak value estimation of bandlimited signals from their samples, noise enhancement, and a local characterization in the neighborhood of an extremum," *IEEE Trans. Signal Processing*, vol. 51, pp. 771–780, Mar. 2003.
- [18] M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "On the peak to average power of OFDM signals based on oversampling," *IEEE Trans. Commun.*, vol. 51, pp. 72–78, Jan. 2003.
- [19] S. Wei, D. Goeckel, and P. Kelly, "A modern extreme value theory approach to calculating the distribution of the peak to average power ratio in OFDM systems," *Proc. IEEE Int. Contr. Conf.*, pp. 1686–1690, 2002.
- [20] R. Motowani and P. Raghavan, *Randomized Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 1995.
- [21] K. Sathananthan and C. Tellambura, "Coding to reduce both PAR and PICR of an OFDM signal," *IEEE Commun. Lett.*, vol. 6, pp. 316–318, Aug. 2002.
- [22] M. Sharif and B. Hassibi, "A deterministic algorithm that achieves PMEPR of  $c \log n$  for multicarrier signals," in *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Processing*, Apr. 2003.



**Masoud Sharif** (S'99) was born in 1977. He received the B.Sc. (with honors) and M.Sc. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1999 and 2001, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, California Institute of Technology, Pasadena.

His research interests are in the areas of information theory, wireless communications, and signal processing. More specifically, he has been working on the capacity of MIMO broadcast channels, peak-to-average power ratio of multicarrier signals, and data collection delay in sensory networks.



**Babak Hassibi** was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1993 and 1996, respectively, all in electrical engineering.

From October 1996 to October 1998, he was a research associate with the Information Systems Laboratory, Stanford University, and from November 1998 to December 2000, he was a Member of the Technical Staff with the Mathematical Sciences Research Center, Bell Laboratories, Murray Hill, NJ. Since January 2001, he has been an assistant professor of electrical engineering at the California Institute of Technology, Pasadena. He has also held short-term appointments at Ricoh California Research Center, Menlo Park, CA, the Indian Institute of Science, Bangalore, India, and Linköping University, Linköping, Sweden. His research interests include wireless communications, robust estimation and control, adaptive signal processing and linear algebra. He is the coauthor of the books *Indefinite Quadratic Estimation and Control: A Unified Approach to  $H^2$  and  $H^\infty$  Theories* (New York: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000). He is a recipient of the Alborz Foundation Fellowship, the 1999 Hugo Schuck Best Paper Award of the American Automatic Control Council, the 2002 National Science Foundation Career Award, the 2002 Okawa Foundation Research Grant for Information and Telecommunications, the 2003 David and Lucille Packard Fellowship for Science and Engineering, and the 2003 Presidential Early Career Award for Scientists and Engineers (PECASE). He was a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY Special Issue on "Space-Time Transmission, Reception, Coding, and Signal Processing" and is currently an Associate Editor Communications for the IEEE TRANSACTIONS ON INFORMATION THEORY.